Foundation of Cryptography (0368-4162-01), Lecture 4

Pseudorandom Functions

Iftach Haitner, Tel Aviv University

November 29, 2011

Section 1

Function Families

function families

1. $\mathcal{F} = \{\mathcal{F}_n\}_{n\in\mathbb{N}}$, where $\mathcal{F}_n = \{f \colon \{0,1\}^{m(n)} \mapsto \{0,1\}^{\ell(n)}\}$
2. We write $\mathcal{F} = \{\mathcal{F}_n \colon \{0,1\}^{m(n)} \mapsto \{0,1\}^{\ell(n)}\}$
3. If $m(n) = \ell(n) = n$, we omit it from the notation
4. We identify function with their description
5. The rv F_n is uniformly distributed over $\mathcal{F}_n$

efficient function families

Definition 1 (efficient function family)

An ensemble of function families $\mathcal{F} = \{\mathcal{F}_n\}_{n\in\mathbb{N}}$ is efficient, if the following hold:

Samplable. $\mathcal{F}$ is samplable in polynomial-time: there exists a PPT that given 1^n, outputs (the description of) a uniform element in $\mathcal{F}_n$.

Efficient. There exists a polynomial-time algorithm that given $x \in \{0,1\}^n$ and (a description of) $f \in \mathcal{F}_n$, outputs $f(x)$.

random functions

Definition 2 (random functions)

For $m, \ell \in \mathbb{N}$, we let $\Pi_{m,\ell}$ consist of all functions from $\{0,1\}^m$ to $\{0,1\}^\ell$.

- It takes $2^m \cdot \ell$ bits to describe an element inside $\Pi_{m,\ell}$.
- We sometimes think of $\pi \in \Pi_{m,\ell}$ as a random string of length $2^m \cdot \ell$.
- $\Pi_n = \Pi_{n,n}$

pseudorandom functions

Definition 3 (pseudorandom functions)

A function family ensemble $\mathcal{F} = \{\mathcal{F}_n : \{0,1\}^{m(n)} \mapsto \{0,1\}^{\ell(n)}\}$ is pseudorandom, if

$$\left|\Pr[\mathsf{D}^{\mathcal{F}_n}(1^n) = 1] - \Pr[\mathsf{D}^{\Pi_{m(n),\ell(n)}}(1^n) = 1\right| = \operatorname{neg}(n),$$

for any oracle-aided PPT D.

1. Suffices to consider $\ell(n) = n$
2. Easy to construct (with no assumption) for $m(n) = \log n$ and $\ell \in \text{poly}$
3. PRF easily imply a PRG
4. Pseudorandom permutations (PRPs)

Section 2

PRF from OWF

the construction

Construction 4

Let $g\colon \{0,1\}^n \mapsto \{0,1\}^{2n}$. Let $g_0(s) = g(s)_{1,\ldots,n}$ and $g_1(s) = g(s)_{n+1,\ldots,2n}$. For s and $x \in \{0,1\}^*$, let f_s be defined as

$$f_s(x) = g_{x_n}(\ldots(g_{x_2}(g_{x_1}(s))))$$

Let $\mathcal{F}_n = \{f_s\colon s \in \{0,1\}^n\}$ and $\mathcal{F} = \{\mathcal{F}_n\}$.

g is efficient function implies that $\mathcal{F}$ is an efficient family.

Theorem 5 (Goldreich-Goldwasser-Micali)

If g is a PRG then $\mathcal{F}$ is a PRF.

Corollary 6

OWFs imply PRFs.

Proof Idea

- Easy to prove for input of length 2.
  Observation: $D = (g(g_0(U_n)), g(g_1(U_n)))$ is pseudorandom:
  Proof: $D' = (g(U_n^{(0)}), g(U_n^1)) \approx_c U_{4n}$ and $D \approx_c D'$.
- Hence we can handle input of length 2
- Extend to longer inputs?
- We show that an efficient sample from the *truth table* of $f \leftarrow \mathcal{F}_n$, is computationally indistinguishable from that of $\pi \leftarrow \Pi_{n,n}$.

Actual proof

Assume $\exists$ PPT D, $p \in$ poly and infinite set $\mathcal{I} \subseteq \mathbb{N}$ with

$$\left|\Pr[\mathsf{D}^{F_n}(1^n) = 1] - \Pr[\mathsf{D}^{\Pi_n}(1^n) = 1]\right| \geq \frac{1}{p(n)}, \tag{1}$$

for any $n \in \mathcal{I}$ and fix $n \in \mathbb{N}$

Let $t = t(n) \in$ poly be a bound on the running time of $\mathsf{D}(1^n)$.

We use D to construct a PPT D' such that

$$\left|\Pr[\mathsf{D}'(U_{2n}^t) = 1] - \Pr[\mathsf{D}'(g(U_n)^t) = 1\right| > \frac{1}{np(n},$$

where $U_{2n}^t = U_{2n}^{(1)}, \ldots, U_{2n}^{(t(n))}$ and
$g(U_n)^t = g(U_n^{(1)}), \ldots, g(U_n^{(t(n))})$.

The hybrid

Let g and f be as in the definition of $\mathcal{F}_n$

Definition 7

For $k \in \{0,\ldots,n\}$, let $\mathcal{H}_k = \{h_\pi \colon \{0,1\}^n \mapsto \{0,1\}^n \colon \pi \in \Pi_{k,n}\}$, where $h_\pi(x) = f_{\pi(x_{1,\ldots,k})}(x_{k+1,\ldots,n})$

- $f_y(\lambda) = y$
- $\Pi_{0,n} = \{0,1\}^n$, and for $\pi \in \Pi_{0,n}$ let $\pi(\lambda) = \pi$

- Note that $\mathcal{H}_0 = \mathcal{F}_n$ and $\mathcal{H}_n = \Pi_{n,n}$
- Can we emulate $\mathcal{H}_k$? We emulate if from D's point of view.
- We present efficient “function family" $\mathcal{O}_k = \{O_k^{s_1,\ldots,s^t}\}$ s.t.
  - $\mathsf{D}^{O_k^{U_{2n}^t}}(1^n) \equiv \mathsf{D}^{H_k}(1^n)$
  - $\mathsf{D}^{O_k^{g(U_n)^t}}(1^n) \equiv \mathsf{D}^{H_{k-1}}(1^n)$

  for any $k \in [n]$, where H_K is uniformly sampled from $\mathcal{H}_k$.

completing the proof

Let $\mathrm{D}'(y)$ return $\mathrm{D}^{O_k^y}(1^n)$ for k uniformly chosen in $[n]$. Hence

$$
\begin{aligned}
&\left|\Pr[\mathrm{D}'(U_{2n}^t = 1]\right| - \Pr[\mathrm{D}'(g(U_n)^t) = 1] \\
&= \left|\sum_{k=1}^{n} \frac{1}{n} \cdot \Pr[\mathrm{D}^{O_k^{U_{2n}^t}}(1^n) = 1] - \sum_{k=1}^{n} \frac{1}{n} \cdot \Pr[\mathrm{D}^{O_k^{g(U_n)^t}}(1^n) = 1]\right| \\
&= \frac{1}{n}\left|\sum_{k=1}^{n} \Pr[\mathrm{D}^{H_k}(1^n) = 1] - \sum_{k=1}^{n} \Pr[\mathrm{D}^{H_{k-1}}(1^n) = 1]\right| \\
&= \frac{1}{n}\left|\Pr[\mathrm{D}^{H_n}(1^n) = 1] - \Pr[\mathrm{D}^{H_0}(1^n) = 1]\right| = \frac{1}{np(n)} \square
\end{aligned}
$$

The family $\mathcal{O}_k$

$$\mathcal{O}_k := \{O_k^{s^1,\ldots,s^t} : s^1,\ldots,s^t \in \{0,1\}^n \times \{0,1\}^n\}.$$

Algorithm 8 ($O_k^{s^1,\ldots,s^t}$)

On the i'th query $x^i \in \{0,1\}^n$:

1. If x^ℓ with $x^\ell_{1,\ldots,k-1} = x^i_{1,\ldots,k-1}$ was previously asked, set $z = s^\ell_{x_k}$ (where ℓ is the minimal such index). Otherwise, set $z = s^i_{x_k}$.
2. Return $f_z(x_{k+1,\ldots,n})$

$\mathcal{O}_k$ is *stateful*.

We need to prove that $\mathsf{D}^{O_k^{U_{2n}^t}}(1^n) \equiv \mathsf{D}^{H_k}(1^n)$ and $\mathsf{D}^{O_k^{g(U_n)^t}}(1^n) \equiv \mathsf{D}^{H_{k-1}}(1^n)$.

$\mathsf{D}^{O_k^{U_{2n}^t}}(1^n) \equiv \mathsf{D}^{H_k}(1^n)$

Proposition 9

For any $\ell, m \in \mathbb{N}$ and any algorithm A, it holds that $A^{\Pi_{\ell,m}} \equiv A^{B_{\ell,m}}$, where the stateful random algorithm $B_{\ell,m}$ answers identical queries with the same answer, and answers new queries with a random string of length m.

Proof? Does the above trivialize the whole issue of PRF?
Let $\widetilde{O}_k$ be the variant that returns z (and not $f_{x_{k+1,\ldots,n}}(z)$) and let $\widetilde{\mathsf{D}}_k$ be the algorithm that implements D using $\widetilde{O}_k$ (by computing $f_{x_{k+1,\ldots,n}}(z)$ by itself).
By Proposition **??**

$$\mathsf{D}^{O_k^{U_{2n}^t}}(1^n) \equiv \widetilde{\mathsf{D}}_k^{\widetilde{O}_k^{U_{2n}^t}}(1^n) \equiv \widetilde{\mathsf{D}}_k^{\pi_{k,n}}(1^n) \equiv \mathsf{D}^{H_k}(1^n) \tag{2}$$

$D^{O_k^{g(U_n)^t}}(1^n) \equiv D^{H_{k-1}}(1^n)$

It holds that

$$D^{O_k^{g(U_n)^t})}(1^n) \equiv D^{O_{k-1}^{U_{2n}^t}}(1^n) \tag{3}$$

Hence, by **??**

$$D^{O_k^{g(U_n)^t}}(1^n) \equiv D^{H_{k-1}}(1^n)$$

Section 3

PRP from PRF

Pseudorandom permutations

Let $\widetilde{\Pi}_n$ be the set of all permutations over $\{0,1\}^n$.

Definition 10 (pseudorandom permutations)

A permutation ensemble $\mathcal{F} = \{\mathcal{F}_n : \{0,1\}^n \mapsto \{0,1\}^n\}$ is a pseudorandom permutation, if

$$\left| \Pr[\mathsf{D}^{\mathcal{F}_n}(1^n) = 1] - \Pr[\mathsf{D}^{\widetilde{\Pi}_n}(1^n) = 1 \right| = \operatorname{neg}(n), \tag{4}$$

for any oracle-aided PPT D

- **??** holds for any PRF

Construction

Construction 11

Given a function family $\mathcal{F} = \{\mathcal{F}_n \colon \{0,1\}^n \mapsto \{0,1\}^n\}$, let $\mathrm{LR}(\mathcal{F}) = \{\mathrm{LR}(\mathcal{F}_n) \colon \{0,1\}2n \mapsto \{0,1\}^{2n}\}$, where $\mathrm{LR}(\mathcal{F}_n) = \{\mathrm{LR}(f) \colon f \in \mathcal{F}_n\}$ and $\mathrm{LR}(f)(\ell, r) = (r, f(r) \oplus \ell)$.
For $i \in \mathbb{N}$, let $\mathrm{LR}^i(\mathcal{F})$ be the i'th iteration of $\mathrm{LR}(\mathcal{F})$.

$\mathrm{LR}(\mathcal{F})$ is always a permutation family, and is efficient if $\mathcal{F}$ is.

Theorem 12 (Luby-Rackoff)

Assuming that $\mathcal{F}$ is a PRF, then $\mathrm{LR}^3(\mathcal{F})$ is a PRP

It suffices to prove the the following holds for any $n \in \mathbb{N}$ (why?)

Claim 13

$|\Pr[\mathrm{D}^{\mathrm{LR}^3(\Pi_n)}(1^n) = 1] - \Pr[\mathrm{D}^{\widetilde{\Pi}_{2n}}(1^n)| = 1] \leq \frac{4 \cdot q^2}{2^n}$,
for *any* q-query algorithm D.

Section 4

Applications

general paradigm

Design a scheme assuming that you have random functions, and the realize them using PRF.

Private-key Encryption

Construction 14 (PRF-based encryption)

Given an (efficient) PRF $\mathcal{F}$, define the encryption scheme $(\mathrm{Gen}, \mathrm{E}, \mathrm{D}))$ se:

Key generation $\mathrm{Gen}(1^n)$ returns $k \leftarrow \mathcal{F}_n$

Encryption $\mathrm{E}_k(m)$ returns $U_n, k(U_n) \oplus m$

Decryption $\mathrm{D}_k(c = (c_1, c_n))$ returns $k(c_1) \oplus c_2$

- Advantages over the PRG based scheme?
- Proof of security